

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
10 May 2001 (10.05.2001)

PCT

(10) International Publication Number  
**WO 01/33437 A1**

(51) International Patent Classification<sup>7</sup>: G06F 17/40, 17/60, 19/00

(21) International Application Number: PCT/US00/41878

(22) International Filing Date:  
3 November 2000 (03.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/163,366 3 November 1999 (03.11.1999) US  
Not furnished 2 November 2000 (02.11.2000) US

(71) Applicant (for all designated States except US):  
PHONE.COM, INC. [US/US]; 800 Chesapeake Drive,  
Redwood City, CA 94063 (US).

(72) Inventor; and

(75) Inventor/Applicant (for US only): WITTEMAN,  
Bradley, James [US/US]; 815 Prospect Row, #7, San  
Mateo, CA 94402 (US).

(74) Agents: WIESNER, Leland, Z. et al.; Fish & Richardson  
P.C., 2200 Sand Hill Road #100, Menlo Park, CA 94025  
(US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

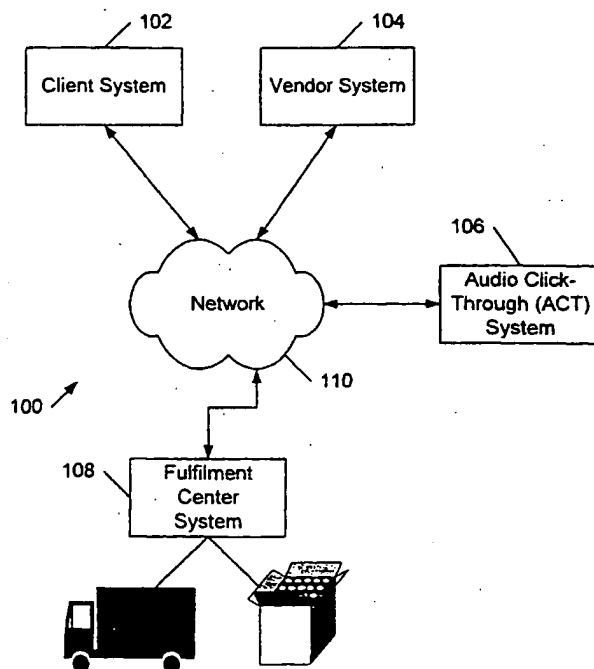
(84) Designated States (regional): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: AUDIO CLICK THROUGH ADVERTISING



(57) Abstract: A system (100) for delivering advertisements over a network includes a vendor system (104) for aggregating multimedia content into an audio click-through (ACT) advertisement (106). The ACT advertisement has an audio portion capable of generating audio output and an interactive portion capable of receiving commands from a user and performing operations in response to the commands. A client system capable of processing the ACT advertisement plays an audio portion of the ACT advertisement and uses an interactive portion of the ACT advertisement with automatic voice recognition to identify a command from the user in response to the audio portion and performs operations in response to the commands.

WO 01/33437 A1



patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

— Before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments.

**Published:**

— With international search report.

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

## AUDIO CLICK THROUGH ADVERTISING.

This application takes priority from U.S. provisional application serial no. 60/163,366, filed November 3, 1999.

### BACKGROUND OF THE INVENTION

5       The present invention relates to computer systems and advertising. In conventional print advertising, images and text provide a two-dimensional experience with limited measurable impact. Print advertisers obtain feedback for an advertisement or ad by displaying the ad and then polling a focus group they believe represents the target market. These polls indirectly show the impact of the ads and effectiveness of advertising costs. Alternatively, print advertisers can also approximate the effectiveness of their ads and return on advertising costs by tracking the phone calls generated through a "Source of Business" tracking code. These codes are typically attached to a telephone number, an address, and/or a coupon along with the print advertisement. In both cases, the feedback and tracking capabilities associated with conventional print ads is limited and at best  
15       imprecise.

      Conventional Internet advertising improves upon traditional print advertising by measuring a "click-through" rate of an ad and its relative success rate compared with other ads. Unlike conventional two-dimensional ads, the Internet allows the advertiser to redirect the audience to websites closely associated with the specific advertisement. This  
20       connects the audience to the advertiser more immediately and capitalizes on advertising expenditures. More importantly, click-through ads allow advertisers to measure effectiveness of certain advertisements and monetize the conversion of these advertisements into purchases.

      Internet advertising unfortunately also has several drawbacks shared by  
25       conventional print advertising. Conventional Internet banner ads do not demand focused attention from the audience. The user can choose to view the banner ads or can completely ignore the banner ad being displayed. Even if the audience does view the banner ad, the banner ad can change so rapidly that the overall advertisement has questionable effectiveness. Consequently the user's attention is split between the  
30       information actually sought, and the advertisements, branding and other information accompanying the information on the page.

Conventional audio ads provided on radio and over the Internet improve upon purely visual ads because the audience can only listen to one audio stream of information at a time. The temporal nature of audio information demands the audience's attention and concentration for a fixed time interval. Unlike visual advertisements, the audience cannot overlook or easily avoid the audio information. This makes audio a more effective medium for delivering advertisements. Examples of audio advertising include radio, television, and audio bulletin board (like those operated by some radio stations to list local events).

Despite these advantages, it remains difficult to determine the value of audio advertising. The impact of conventional audio advertisements cannot be measured readily and connected with specific commercial transactions. The audience may listen to an advertisement for a product and decide at a later time not to purchase the product. Others who purchase an advertised product or service may not have heard the audio advertisements. Essentially, no direct connection can be made between audio ads and commercial transactions.

#### SUMMARY OF THE INVENTION

One aspect of the present invention includes system for delivering an audio click-through (ACT) advertisement over a network. This system includes a vendor system for aggregating multimedia content into an ACT advertisement wherein the ACT advertisement has an audio portion capable of generating audio output and an interactive portion capable of receiving a command from a user and performing operations in response to the command. The system further includes a client system that plays the audio portion of the ACT advertisement and processes the interactive portion with automatic voice recognition to identify a command from the user in response to the audio portion. The client system then performs an operation in response to the command.

Another aspect of the present invention includes just a vendor system. The vendor system aggregates multimedia content into an ACT advertisement. The ACT advertisement has an audio portion capable of generating audio output and an interactive portion capable of receiving a command from a user and performing an operation in response to the command.

Yet another aspect of the invention includes just a client system. The client system plays an audio portion of an ACT advertisement and processes an interactive portion of the ACT advertisement with automatic voice recognition. The automatic voice

recognition identifies a command from the user and executes an operation in response to the command.

Yet another aspect of the invention includes an ACT system. The ACT system stores and forwards ACT advertisements having an audio portion capable of generating audio output and an interactive portion capable of receiving commands from a user and performing operations in response to the commands. The ACT system selects the ACT advertisement according to a selection criteria transmitted from a client system capable of processing the ACT advertisement.

Aspects of the invention further include a method of creating an ACT advertisement using multimedia. The method receives a multimedia database having a collection of audio clips, sequences the audio clips into an audio portion having one or more time intervals, associates one or more audio commands with one or more time intervals, and assigns an operation to the one or more audio command thus creating an interactive portion. Combining the audio portion with the interactive portion creates the ACT advertisement.

In addition, aspects of the invention include a method of processing an ACT advertisement. The ACT advertisement has an audio portion capable of generating audio output and an interactive portion capable of receiving an audio command from a user and performing operations in response to the audio command. Processing the ACT advertisement includes presenting the ACT advertisement to the user during a time interval, receiving an audio command from the user during the time interval and, determining if the user desires receiving additional information based upon the audio command received. One or more predetermined operations are performed in response to the command wherein the operation includes generating additional audio output.

Yet another aspect of the invention includes a method of delivering advertisements over a network by receiving a request for an audio click-through (ACT) advertisement wherein the ACT advertisement has an audio portion capable of generating audio output and an interactive portion capable of receiving an audio command from a user and performing operations in response to the audio command. receiving user profile information associated with the request, selecting an ACT advertisement from one or more ACT advertisements that corresponds to the user profile information, and delivering the selected ACT advertisement to a client system capable of processing the ACT advertisement.

Aspects of the invention also include a method of using advertisement to generate revenue. This method includes receiving a request for an audio click-through (ACT) advertisement wherein the ACT advertisement has an audio portion capable of generating audio output and an interactive portion capable of receiving an audio command from a user and performing operations in response to the audio command, delivering an ACT advertisement in response to the request for the ACT advertisement, and charging a fee when the ACT advertisement is delivered.

Advantages associated with implementations of the invention include one or more of the following. Users can interact with multimedia advertisements by responding directly to an advertisement using audio commands. Essentially, the users provide verbal commands to "click-through" an advertisement. This interactivity enables an advertiser to provide more targeted information while tracking and measuring the effectiveness of an advertisement. The advertiser capitalizes on an advertisement's impression by providing the user with additional information or a commercial transaction tailored to the user's request. An intuitive speech interface makes audio click-through easier to use than Internet browsers and other computer-based systems using conventional keyboard and mouse driven solutions.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features of the invention will become apparent from the description, the drawings, and the claims.

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is block diagram of a network having audio click-through (ACT) capabilities.

FIG. 2 illustrates an example client system having multimedia capabilities and used to prevent ACT advertisements.

FIG. 3 is a block diagram of a vendor system used to create ACT advertisements.

FIG. 4 is a block diagram of an ACT system that processes ACT advertisements.

FIG. 5 is a flowchart diagram of the operations used to create an ACT advertisement on the vendor system.

FIG. 6 is a flowchart diagram of the operations for delivering ACT advertisements from an ACT system to a client system.

FIG. 7 is a flowchart of the interactions associated with an ACT ad.

FIG. 8 is a flowchart of the interactions associated with an ACT ad having a transaction.

### DETAILED DESCRIPTION

FIG. 1 is block diagram of a network 110 having audio-click-through (ACT) capabilities. Network 110 includes a client system 102, a vendor system 104, an audio click-through (ACT) system 106, and fulfillment centers 108 connected together over a network 110. These systems are separated to facilitate scalability and modular systems development practices. Likewise, the specific function each system actually performs depends on the processor capabilities, storage capacity of each system, and network bandwidth between each of the systems on network 110. With different hardware capacities, the functions might be distributed differently and could even be implemented on a single computer system.

Client system 102 is a computer-based platform that processes and delivers ACT ads having multimedia information and voice interactive capabilities. For example, client system 102 can be a cellular phone or part of a car stereo having voice synthesis, voice recognition, and processing capabilities to process these subsystems. The ACT ad provides audio information, receives audio commands from a user, and processes audio commands using automatic voice recognition/synthesis.

Vendor system 104 is another computer-based platform used to aggregate content and create ACT ads for delivery over network 110. For example, manufacturers use vendor system 104 to assemble ACT ads for their particular manufactured products. Similarly, advertisement agencies use vendor system 104 to mix and match multimedia content to create many different ads for their client's products and services. ACT ads can also deliver detailed information on other subjects including news, sports, weather and other information.

Once the ACT ads are created, ACT system 106 processes and delivers the ads to client system 102 connected to network 110. Processing ACT ads involves selecting a proper ad and verifying that users can interact with the ad on client system 102. For example, the language and currency must match the users desired language and currency for entering into transactions. The code must be compatible with the hardware parameters of client system 102 and must also conform to the user-interface (i.e., buttons, graphical user interface, screen size, voice recognition capabilities) of client system 102.

If a transaction is associated with an ACT ad, a fulfillment center system 108 receives and fulfills requests from the ad. For example, an interaction with an ACT ad may result in a request to mail additional information or purchase a product or service.

In operation, systems in FIG. 1 work together to provide ACT ads while a user is  
5 listening to an audio stream of information. Vendor system 104 is used to create an ACT ad and determine how audio commands interact with the ACT ad. The ACT ad is uploaded to ACT system 106 where it is stored before being delivered to client system 102.

A user on client system 102 plays an audio stream developed by a content  
10 provider. The content provider designs the audio stream to be filed with ACT ads at predetermined time intervals to generate advertising revenue. As each time interval is encountered, client system 102 transmits a request for an ACT ad on behalf of the particular audio stream being played. In one implementation, a "cookie" with a user's profile along with contextual or substantive information describing the contents of the  
15 audio stream being played is sent to ACT system 106. ACT system 106 uses the user profile information and the contextual information about the audio stream to select an appropriate ACT ad. The selected ACT ad is uploaded to client system 102 and played for the user at the particular time interval in audio stream. The user interacts with the audio portion of the ACT ad using one or more audio commands. ACT system 106 sends  
20 operations to client system 102 in response to the audio commands. Executing the operations on client system 102 provides additional information to the user or initiates a transaction to purchase the goods or services.

Referring to FIG. 2, the block diagram illustrates one example of client system 102 used to process and present ACT ads to a user. In this example, client system 102  
25 has multimedia capabilities including an audio input 204, a visual display 206, an audio output 208, selectable areas 210, and an antenna 212.

Audio input 204 receives voice commands from a user wishing to interact with an ACT ad. The voice commands processed by a voice recognition process or subsystem in client system 102 identifies which operations in the ACT ad should be performed.  
30 Alternatively, voice recognition processing can be done remotely on ACT system 106 with the results transmitted to client system 102 over network 110. Details on using audio commands with an ACT ad are described in further detail below.

Visual display 206 provides a display area where the user can see information on a product or service. The display area shows the product being offered in the ad or a



demonstration of the service being offered. It also displays a set of voice commands for interacting with the ACT ad. For example, as the ad is running the display can display visual cues of the commands to be spoken such as "Purchase product", "More information", "Mail discount coupons to home address", or "Exit ad". These visual cues  
5 spell-out commands in text on the display or provide icons for the commands using recognizable or universal symbols.

Audio output 208 provides the audio associated with the ACT ad and the audio stream. The audio stream is interspersed with one or more ACT ads. A user can interact with each ACT ad using a predetermined set of audio commands. Acceptable commands  
10 used in the ACT ad are determined when the ACT ad is created on vendor system 104. Different user commands cause different selections to be offered aurally through audio output 208. For example, a user providing a "purchase" command is offered different payment options such as credit, ATM, or electronic transfer of funds while a user providing a "more information" command is offered additional information and the  
15 option of providing a mailing address or email address for the delivery.

Selectable areas 210 are an alternate method of interacting with the ACT ad. This is useful when noise levels would interfere with audio input 204 and speech recognition as well as in cases where the user is mute or speaks a different language. Accordingly, each of the selectable areas 210 can have predetermined commands associated with the  
20 selections or can be dynamically assigned via a vis visual display 206. For example, in processing an ACT ad visual display 206 is programmed to show the selectable areas 210 having commands such as "Purchase product", "More information", "Mail discount coupons to home address", or "Exit ad". Both the ACT ad and selectable areas 210 are customized to the users native language preferences.

25 Antenna 212 provides connectivity to network 110 and facilitates delivery of ACT ads to client system 102 as well as returning any commands from the user to ACT system 106 or vendor system 104. Depending on the specific configuration, antenna 212 can be a wireless local area network connection using the TCP/IP protocol, a two-way radio transmission protocol, a wireless communication protocol such as CDMA, GSM, TDMA,  
30 or AMPS, or any other way to provide connectivity and communication to client system 102. Alternatively, antenna 212 can be replaced with a non-wireless connection to network 110 by way of a cable, DSL, or telephone connection.

FIG. 3 is a block diagram of vendor system 104 used to create ACT ads. A computer-based implementation of vendor system 104 includes a memory 302, a

processor 304, a communication port 306, a secondary storage 308, a multimedia database 310, a transaction database 312, and input-output ports 314, connected over a network or bus 316.

Processor 304 can be a general purpose processor such as a Pentium processor by Intel, Inc., a more specialized processor such as an ARM processor by ARM Limited, or an embedded processor such as one of the MIPs line of processors designed by MIPs, Inc. Communication port 306 provides network connectivity to network 110 using a network protocol such as TCP/IP over an Ethernet-type medium. Alternatively, communication port 306 accommodates wireless communications described above and the corresponding air interfaces that they require.

Many different types of information used by vendor system 104 are stored in secondary storage 308. For example multimedia database 310 stores clips of sound, video, and images. To save space, entries in multimedia database 310 may be shared by one or more ACT ads. If storage capacity is plentiful then multimedia information from multiple ACT ads are stored separately.

Transaction database 312 holds information on electronic commerce transactions that occur when a user interacts with an ACT ad. This information is typically gathered remotely through ACT system 106 and then forwarded to the proper vendor system 104 and/or fulfillment center system 108 for further processing. Information stored in transaction database 312 includes a user's previous orders, preferences, credit card numbers and addresses. An automatic number identification (ANI) cross-references entries in transaction database 312 and facilitates rapid access of information when the user calls up on a telephone or telephone-device compatible with ANI-type identification. Automatic location information (ALI) can also be used to cross-reference entries in transaction database 312 and influence details of a transaction. For example, an ACT ad for a restaurant in a chain of restaurants can use the ALI to locate the take-out food restaurant closest to the user and then place the users order.

A collection of software modules in memory 302 on vendor system 104 facilitates developing and processing of ACT ads. ACT multimedia source module 318 provides an interface to create a framework for the ACT ad. This includes identifying the location in time of audio segments for advertising and the location of the interactive audio portions of the ad. Additionally, ACT multimedia source module 318 provides hooks to synchronize visual information such as images and video with audio in the ACT ad. For example, an ACT ad can include a 30 second advertisement, a set of 10 predetermined

commands that perform different operations and have hooks for displaying images at certain time intervals. This framework can be described using XML or other types of markup languages and compiled using existing XML compilers modified to process this type of information.

5           ACT data module 320 fills the framework of the ACT ad with content and interactive information. Essentially, ACT data module 320 interfaces with content on multimedia database 310 and transaction formats on transaction database 312. This includes providing audio, video, and image content for the advertisements and content for the interactive audio portions. This organization allows for different languages and  
10           currencies to use the same ad framework depending on a users language and currency preference. For example, a French ACT ad uses the same framework as an English ACT ad by specifying the language as French and the currency as Francs rather than British pounds.

          ACT generator module 322 parses and combines information in ACT data module  
15           320 with ACT multimedia source module 318 and creates the ACT ad for storage in ACT Ad module 324. If some aspect of the information in ACT data module 320 or ACT multimedia source module 318 is not properly formed, ACT generator module 322 flags the mistake for correction.

          ACT ad module 324 compresses the size of ACT ad 324 by eliminating redundant  
20           copies of multimedia in an ACT ad as well as performing run-length encoding and other bitwise algorithmic compression techniques on the actual ACT ad 324. Compression is important to reduce both storage requirements and bandwidth requirements for sending the ACT ads.

          Private transaction information in transaction database 312 is accessed through  
25           ACT ad transaction module 326. Typically, private transaction information includes any information a vendor would like to see before approving a transaction. For example, private transaction information can include a person's credit report, birthdate, social security number, credit-card numbers and home address. Vendor system 104 keeps track of this information to better serve their client base while ensuring the transaction  
30           information is kept secure and confidential on vendor system 104.

          Runtime environment 328 manages the allocation and usage of resources on vender system 104. This can be implemented using general purpose operating systems such as UNIX, Linux, and Windows or real-time operating systems for faster responsiveness.

FIG. 4 is a block diagram illustration of ACT system 106 for processing ACT ads. A computer-implementation of ACT system 106 includes a memory 402, a processor 404, a communication port 406, a secondary storage 408, a keyword database 410, a multimedia advertisement database 412, a multimedia output database 414, input-output  
5 ports 416, connected together over a network or bus 418. Elements named in FIG. 4 operate in a similar manner as the corresponding like-named elements in FIG. 3 described above. Unlike vendor system 104 described in FIG. 3, ACT system 106 has different databases associated with secondary storage 408 and different modules in memory 402.

Keyword database 410 in secondary storage 408 includes words used as selections  
10 in one or more of the ACT ads. These words are compiled, formatted, indexed and then stored in keyword database 410 for rapid retrieval. Any additional information needed for voice recognition and voice synthesis of these words is also included in keyword database 410.

Multimedia advertisement database 412 includes all the multimedia information  
15 included in the one or more ACT ads. For example, this includes images, videos, and audio clips created on vendor system 104 and uploaded over network 110 to ACT system 106. Sequencing and timing information is associated with the multimedia information and then stored in multimedia advertisement database 412.

The operations performed by the ACT ad are kept in multimedia output database  
20 414. Each word in keyword database 410 is associated with a particular operation in multimedia output database 414. When a keyword in keyword database 410 is recognized, a corresponding operation in output database 414 is performed. For example, if a keyword phrase "send me a voucher" is identified, a set of predetermined operations stored in output database 414 forwards the voucher automatically to the user's home  
25 address. Detecting other keywords such as "give me all options" generates voice synthesis of words on client system 102 listing the various options available for interacting with the ACT ad.

Modules in memory 402 are for processing and delivering ACT ads from ACT  
system 106 to users of client system 102. Each user provides preferences and interests  
30 stored using user profile database API 420. The information is gathered through a voluntary registration process which can occur either using voice recognition or by entering data on a website using a keyboard and mouse. The preferences include details on how a user likes information to be presented and formatted. For example, a user can

specify the language and currency to present audio information and perform electronic transactions.

User profile database API 420 facilitates transmission of other information about a person's mailing address, transaction history, product and service interests, age, sex, and other useful demographic information. Some additional information about the user is inferred using automatic location identification (ALI) retrieved using the users phone number and/or GPS and an automatic number identification (ANI) of the users phone number.

Information in user profile database helps create targeted marketing campaigns that are efficient and beneficial to the consumer as well as the vendor. For example, ACT system 106 uses user profile database API 420 to identify which of the many advertisements would most interest the user. To keep this information private, personal information is stored on a user's computer in the form of "cookies" and accessed as necessary through user profile database API 420 using a secure communication mechanism such as SSL or public-key encryption.

Keyword database API 422, multimedia advertisement API 424, and multimedia output database API 426 are all interfaces to the like-named databases previously described. These interfaces provide ACT ads with database access as the ACT ads are processed.

For example, an ACT ad retrieves multimedia content from multimedia advertisement database 412 using multimedia advertisement database API 422. The content is prepared in advance on different vendor system 104 as described above and uploaded to ACT system 106 over network 110. The specific multimedia content retrieved depends on the user's response and the operations described in multimedia output database 414 accessed through multimedia output database API 426. For example, if the user requests additional information, more content is provided with the information. Alternatively, when a user expresses a desire to purchase a product or service, the ACT ad enters into a transaction according to operations stored in multimedia output database 414.

Automatic speech recognition/synthesis component 428 access keyword database 410 through keyword database API 422 to determine if a valid command has been received. Entries in keyword database 410 are compared with each command to determine what actions to perform. For example, automatic speech recognition/synthesis component 428 compares words and phrases such as "send me a voucher", "purchase one

now", or "not interested" until a match is found. Nuance Communications of Menlo Park, California provides one type of automatic speech recognition/synthesis component 428 capable of being used in this application.

Information on transactions and user selections are tracked using advertisement tracking agent 430. Advertisement tracking agent 430 works with ANI and ALI information and other software logic to determine location information and phone numbers used by the user. Also, advertisement tracking agent 430 uses heuristics to categorize a users particular interests over time. For example, advertisement tracking agent 430 identifies each ACT ad a user interacts with to determine which ads are most likely to appeal to the user in the future. This involves detecting trends in the type of information being requested and the content being provided. In some cases, shorter ads may be determined to be more effective than longer ads. Similarly, ads providing more detail may be more effective for expensive items while ads with less detail may be determined to be more effective for lower cost items.

Referring to FIG. 5, a flowchart diagram provides the operations used to create an ACT ad using vendor system 104. Initially, an ad designer creates a multimedia database for audio click-through ads (502). This involves defining the database schema to hold the information and then populating the database with image, video, audio, and other types of multimedia information. Information in the multimedia database is indexed for fast retrieval and cross-referencing with other database tables. Multimedia database 310 in FIG. 3 holds the multimedia database information for creating one or more particular ACT ads.

The ad designer then designs the interactions to occur in the ACT ad and which multimedia to use in the ad (504). This involves selecting preexisting audio, video and images from the multimedia database and in some cases recording audio clips. Each of these different media types are sequenced and associated with one or more audio commands. For example, the ad designer may display a list of available audio commands in a display portion of the ACT ad on the display of client device 102. This list of available audio commands provides a visual cue of the audio commands available for the user to interact with the ACT ad during playback. The ad designer also associates one or more audio commands with time intervals in the sequence of audio clips. These commands typically relate to the information provided during the particular time interval. Example operations include playing additional audio information and engaging in a

transaction. Multimedia source module 318 in FIG. 3 holds sequencing and interactive information for the ACT ad.

ACT generator module 322 uses the multimedia information in ACT data module 320 and ad design ACT multimedia source module 318 to generate an ACT ad for a specific target device (506). Information in both ACT data module 320 and ACT multimedia source module 318 are parsed and errors that may exist are indicated. These errors in formatting, content, and organization are corrected and eventually the ACT ad is created. The ad designer then connects one or more events in the ACT ad with corresponding transactional processes (508). For example, a user selecting a predetermined keyword such as "purchase" during a predetermined time interval in the ACT ad will automatically enter into a transaction process to purchase the advertised goods or services. Once the ACT ad is complete, it is uploaded over the network to a delivery system (510) such as ACT system 106 that delivers the ACT ad to client system 102.

FIG. 6 is a flowchart of the operations that ACT system 106 uses to deliver ACT ads. ACT ads are initially loaded onto ACT system 106 over network by one or more vendor systems 104 (602). Each ACT ad contains a different advertisement and can be targeted to a specific target device such as a cellular phone or two-way car radio. ACT system 106 receives a request for an ACT ad from client system 102 (604). Along with the request, ACT system 106 receives a "cookie" from client system 102 providing user profile information. In an alternate implementation, user profile information can be derived from demographic information and account data for a user found in a carrier's database or a vendor database rather than information in a "cookie". Further the ACT ad can be provided to client system 102 without a request and instead by pushing the ACT ad. Additionally, ACT system 106 receives contextual information describing the substantive audio information being delivered to a user. The contextual or substantive information describes the kind of audio information being provided. For example, the audio information can be news, music, weather, or sports information. In alternative

ACT system 106 selects an appropriate ACT ad to deliver to client system 102 based on user profile information and the contextual information (606). For example, if a user enjoys expensive luxury items and is listening to audio information about automobiles then an ACT ad for Mercedes Benz is delivered to client system 102. If the audio information is about watches, then an ACT ad for Rolex Watches would be delivered to client system 102 instead.

ACT system 106 processes the user's audio command responses to the ACT ad (608). For example, client system 102 transmits responses from the user over network 110 to ACT system 106 where the command is processed using speech processing techniques. ACT system 106 identifies the command and sends back the operation(s) for client system 102 to present (610). For example, ACT system 106 may provide additional audio clips for client system 102 to play or may engage in a transaction to purchase a service or product. Each time ACT system 106 delivers an ACT ad to client system 102, the company providing the product, service, or information is charged a transaction fee (612). The company can also be charged additional fees for locating a lead for a transaction, processing a transaction, or merely providing additional information on the ACT ad to client system 102. In an alternate implementation, a percentage of the transaction can be credited to a carrier transmitting the information and/or the vendor involved in performing the transaction with the user.

FIG. 7 is a flowchart of the operations associated with an ACT ad. The ACT ad is initially presented to a user on client system 102 (702). The ACT ad can be downloaded onto client system 102 or streamed over network 110 to client system 102. Client system 102 can be wireless telephone device, a wired telephone device, a personal digital assistant (PDA), a radio device having a microphone, or any other type of device capable of receiving audio information, providing audio information, and optionally displaying images or video.

The user receives audio information from client system 102 along with a number of commands for interacting with the ACT ad (704). For example, commands for interacting with the ACT ad can be made part of the audio whereby each of the different available commands are voice synthesized or announced with a predetermined voice recording. The commands for interacting with the ad also can be displayed on a display portion of client system 102 if such as display exists and is available.

The user speaks audio commands while the ACT ad is presented. The client system 102 determines if the user is requesting additional information with the audio commands (706) or is not interested in interacting with the ACT ad. If additional information is requested, client system 102 receives and processes commands (708). In one implementation, client system 102 transmits commands to ACT system 106 to identify further operations. Alternatively, client system 102 is equipped with automatic speech recognition/synthesis module 428, keyword database 410, and multimedia output database 414 and is capable of processing the user's commands locally.



Client system 102 performs the predetermined operations in response to the audio commands provided (710). This includes presenting audio that describes a product or service in further detail, interacting with the user using a set of questions, gathering information from the user, or providing real-time information such as a price quote on a stock, commodity or other product. The user's responses to the predetermined operations are then provided to vendor system 104 (712). ACT system 106 uses the user's responses to compile demographic statistics and charge a transaction fee to the company providing the advertised product or service (714).

When the user indicates no interest in receiving more information from the ACT ad (706), the remainder of the audio portion of the ACT ad is presented on client system 102 (716). The substantive information continues being provided and another ACT ad is queued for presentation at the next time slot designated for advertising (718).

FIG. 8 is a flowchart of the operations for entering into a transaction through an ACT ad. The ACT ad is initially presented to a user on client system 102 (802). The ACT ad can be downloaded onto client system 102 or streamed over network 110 to client system 102. The user receives audio information from client system 102 along with a number of commands for interacting with the ACT ad (804). For example, commands for interacting with the ACT ad are made part of the audio whereby each of the different available commands are voice synthesized or announced with a predetermined voice recording. The commands for interacting with the ad also can also be displayed on a display portion of client system 102 if such as display exists and is available.

The user speaks audio commands while the ACT ad is presented. The client system 102 determines whether or not the user wishes to enter into a transaction by analyzing the audio commands (806). If the user wants to begin a transaction, client system 102 uses the ACT ad to obtain transaction details from the user (808). For example, the ACT ad may request credit card and billing address information from the user. Client system 102 transmits the commands and transaction information to ACT system 106 where the commands are identified and the transaction information is stored for future reference. Alternatively, client system 102 is equipped with automatic speech recognition/synthesis module 428, keyword database 410, and multimedia output database 414 and is capable of processing the users commands and transaction information locally.

Client system 102 performs a predetermined set of operations associated with the transaction (810). This includes presenting interactive audio information that describes a product or service in further detail, gathering other transaction details from the user, performing credit verification if necessary, and may even include checking for available  
5 inventory. Additional transaction information gathered from the user in response to these predetermined operations is then provided to vendor system 104 for further processing (812). ACT system 106 also uses the transaction information to compile demographic statistics and charge a transaction fee to the company providing the product or service (814).

10 When the user indicates no interest in entering into a transaction (806), the remainder of the audio portion of the ACT ad is presented on client system 102 (816). The substantive information continues being provided and another ACT ad is queued for presentation at the next time slot in the audio stream designated for advertising (818).

The invention can be implemented in digital electronic circuitry, or in computer  
15 hardware, firmware, software, or in combinations of them. Apparatus of the invention can be implemented in a computer program product tangibly embodied in a machine-readable storage device for execution by a programmable processor; and method steps of the invention can be performed by a programmable processor executing a program of instructions to perform functions of the invention by operating on input data and  
20 generating output. The invention can be implemented advantageously in one or more computer programs that are executable on a programmable system including at least one programmable processor coupled to receive data and instructions from, and to transmit data and instructions to, a data storage system, at least one input device, and at least one output device. Each computer program can be implemented in a high-level procedural or  
25 object-oriented programming language, or in assembly or machine language if desired; and in any case, the language can be a compiled or interpreted language. Suitable processors include, by way of example, both general and special purpose microprocessors. Generally, a processor will receive instructions and data from a read-only memory and/or a random access memory. Generally, a computer will include one or  
30 more mass storage devices for storing data files; such devices include magnetic disks, such as internal hard disks and removable disks; magneto-optical disks; and optical disks. Storage devices suitable for tangibly embodying computer program instructions and data include all forms of non-volatile memory, including by way of example semiconductor memory devices, such as EPROM, EEPROM, and flash memory devices; magnetic disks

such as internal hard disks and removable disks; magneto-optical disks; and CD-ROM disks. Any of the foregoing can be supplemented by, or incorporated in, ASICs (application-specific integrated circuits).

- The invention has been described in terms of particular embodiments. Other
- 5   embodiments are within the scope of the following claims. For example, the steps of the invention can be performed in a different order and still achieve desirable results.

## WHAT IS CLAIMED IS:

1. A system for delivering advertisements over a network comprising:  
a vendor system for aggregating multimedia content into an audio click-through (ACT) advertisement wherein the ACT advertisement has an audio portion  
5 capable of generating audio output and an interactive portion capable of receiving commands from a user and performing operations in response to the commands; and  
a client system capable of processing the ACT advertisement by playing an audio portion of the ACT advertisement, using the interactive portion with automatic voice recognition to identify a command from the user in response to the audio portion  
10 and performing an operation in response to the command.
2. The system of claim 1, further comprising:  
an audio click-through system that stores the ACT advertisement and transmits the ACT advertisement to the client system according to a selection criteria.
3. The system of claim 2, wherein the selection criteria depend on  
15 personalized information in a user profile.
4. The system of claim 1, further comprising:  
a fulfillment center system that processes a transaction in furtherance of the delivery of products or services in response to the command from the user.
5. The system of claim 1, wherein the ACT advertisement further includes a  
20 display portion capable of being displayed on a display of the client system wherein the display portion includes visual information on a product or service being advertised and a list of commands for interacting with the ACT advertisement.
6. The system of claim 1, wherein the vendor system aggregates multimedia information into an ACT advertisement describing a product.
- 25 7. The system of claim 1, wherein the vendor system aggregates multimedia information into an ACT ad describing a service.
8. The system of claim 1, wherein the client system is selected from a set of devices including a cellular phone, a two-way radio, and a wireless personal digital assistant.

9. The system of claim 1, wherein the vendor system aggregates multimedia information into an ACT ad describing a news event.

10. A vendor system for aggregating multimedia content into an audio click-through (ACT) advertisement having an audio portion capable of generating audio output  
5 and an interactive portion capable of receiving a command from a user and performing an operation in response to the command.

11. The vendor system in claim 10, wherein the ACT advertisement further includes a display portion for providing visual information on a product or service being advertised and a list of commands for interacting with the ACT advertisement.

10 12. A client system capable of processing an ACT advertisement having an audio portion and an interactive portion, by playing the audio portion of the ACT advertisement and processing the interactive portion of the ACT advertisement with automatic voice recognition that identifies a command from a user in response to the audio portion and executes an operation in response to the command.

15 13. An audio click-through system that stores and forwards an ACT advertisement having an audio portion capable of generating audio output and an interactive portion capable of receiving a command from a user, wherein the ACT advertisement performs an operation in response to the command and wherein the ACT advertisement is selected according to a selection criteria is transmitted to a client system  
20 capable of processing the ACT advertisement.

14. A method of creating an interactive advertisement using multimedia, comprising:

receiving a multimedia database having a collection of audio clips;  
sequencing the audio clips in an audio portion divided into a set of time  
25 intervals;

associating an audio command with a time interval in the set of time intervals in the sequence of audio clips, wherein the audio command relates to information provided during the time interval;

assigning an operation to the audio command creating an interactive  
30 portion, wherein the operation is performed in response to receiving the audio command;  
and

combining the audio portion and the interactive portion to create an audio click-through (ACT) advertisement.

15. The method of claim 14, wherein the operation performed in response to receiving the audio command includes performing a transaction.

5        16. The method of claim 14, wherein the ACT advertisement is created for processing by a specific target device.

17. The method of claim 14, wherein the multimedia database includes visual information and the visual information includes information on a product or service being advertised and a list of commands for interacting with the ACT advertisement.

10        18. A method of processing an interactive advertisement, comprising:  
receiving an audio click-through (ACT) advertisement having an audio portion capable of generating audio output and an interactive portion capable of receiving an audio command from a user and performing an operation in response to the audio command;

15        presenting the ACT advertisement to the user during a time interval;  
receiving an audio command from the user during the time interval and in response to the ACT advertisement presented; and  
performing one or more predetermined operations in response to the audio command received.

20        19. The method of claim 18, wherein the ACT advertisement presented to the user is determined according to information in a user profile.

20. The method of claim 19, wherein the user profile includes demographic information describing the user.

21. The method of claim 19, wherein the user profile includes information on  
25 one or more previous selections made by the user.

22. The method of claim 18, wherein presenting the ACT advertisement includes playing the audio portion of the ACT advertisement to generate audio output over a time interval.

23. The method of claim 18, wherein presenting the ACT advertisement further includes displaying a display portion of the ACT advertisement having visual information on a product or service being advertised and a list of commands for interacting with the ACT advertisement.

5           24. The method of claim 18, wherein receiving the audio command includes processing the audio command using automatic speech recognition.

25. The method of claim 24, wherein processing the audio command using automatic speech recognition further comprises comparing the audio command against a set of keywords, wherein each keyword is associated with an operation.

10           26. The method of claim 25, wherein an operation is selected from a set of operations including providing audio output, displaying information on a display, and entering into a transaction.

27. The method of claim 26, wherein the information displayed on a display is selected from a set of displayable items including a web page, a facsimile message, an electronic mail, a short message system (SMS), and a wireless application protocol (WAP) item.

28. The method of claim 18, further comprising,  
charging a transaction fee when the one or more predetermined operations are performed.

20           29. A method of delivering advertisements over a network comprising:  
receiving a request for an audio click-through (ACT) advertisement  
wherein the ACT advertisement has an audio portion capable of generating audio output  
and an interactive portion capable of receiving an audio command from a user and  
performing an operation in response to the audio command;  
25           receiving user profile information associated with the request;  
selecting an ACT advertisement that corresponds to the user profile  
information; and  
delivering the selected ACT advertisement to a client system capable of  
processing the ACT advertisement.

30           30. The method of claim 29, further comprising,

receiving a command in response to the selected ACT advertisement;  
transmitting an operation to the client system in response to the received  
command. .

31. The method of claim 30, further comprising:  
5 charging a fee for the selected ACT advertisement. .
32. The method of claim 29, wherein the user profile information is contained  
in a "cookie".
33. The method of claim 29, wherein selecting an ACT advertisement further  
includes comparing contextual information provided by the client system describing an  
10 audio stream being played with information describing the ACT advertisement  
advertisements.
34. The method of claim 29, wherein delivering the selected ACT  
advertisement to the client system includes transmitting the selected ACT advertisement  
over a network.
- 15 35. A method of using an advertisement to generate revenue, comprising:  
receiving an audio click-through (ACT) advertisement wherein the ACT  
advertisement has an audio portion capable of generating audio output and an interactive  
portion capable of receiving an audio command from a user and performing a operation in  
response to the audio command;  
20 delivering the ACT advertisement to the user; and  
charging a fee for the ACT advertisement.
36. The method of claim 35, wherein the ACT advertisement is selected from  
one or more ACT advertisement based on information in a user profile.
37. The method of claim 35, wherein the ACT advertisement is selected from  
25 one or more ACT advertisement based on information describing information in an audio  
stream.
38. A computer program product, tangibly stored on a computer-readable  
medium, for delivering advertisements over a network, comprising instructions operable  
to cause a programmable processor to:



receive a request for an audio click-through (ACT) advertisement wherein the ACT advertisement has an audio portion capable of generating audio output and an interactive portion capable of receiving an audio command from a user and performing an operation in response to the audio command;

- 5                   receive user profile information associated with the request;  
                  select an ACT advertisement that corresponds to the user profile information; and  
                  deliver the selected ACT advertisement to a client system capable of processing the ACT advertisement.

10

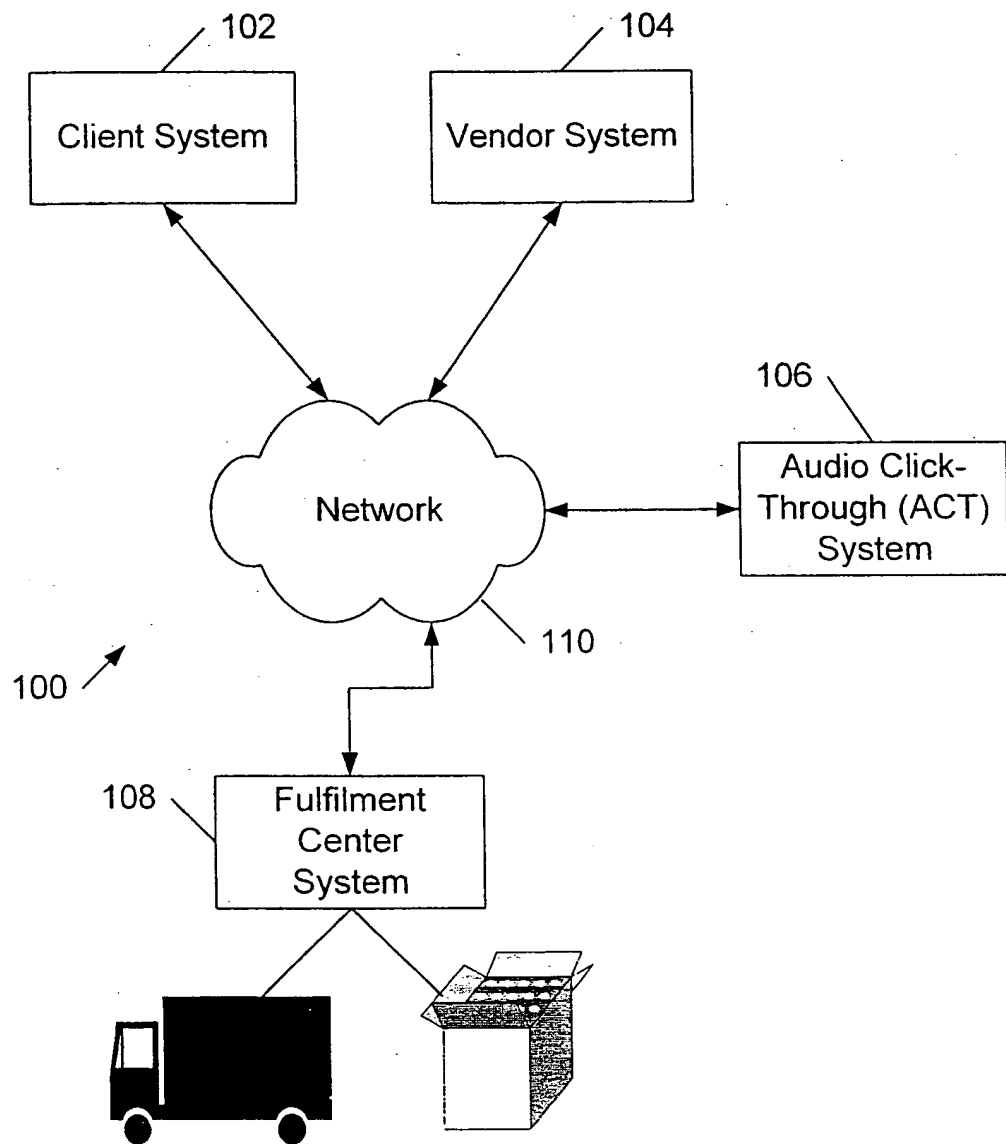


FIG 1

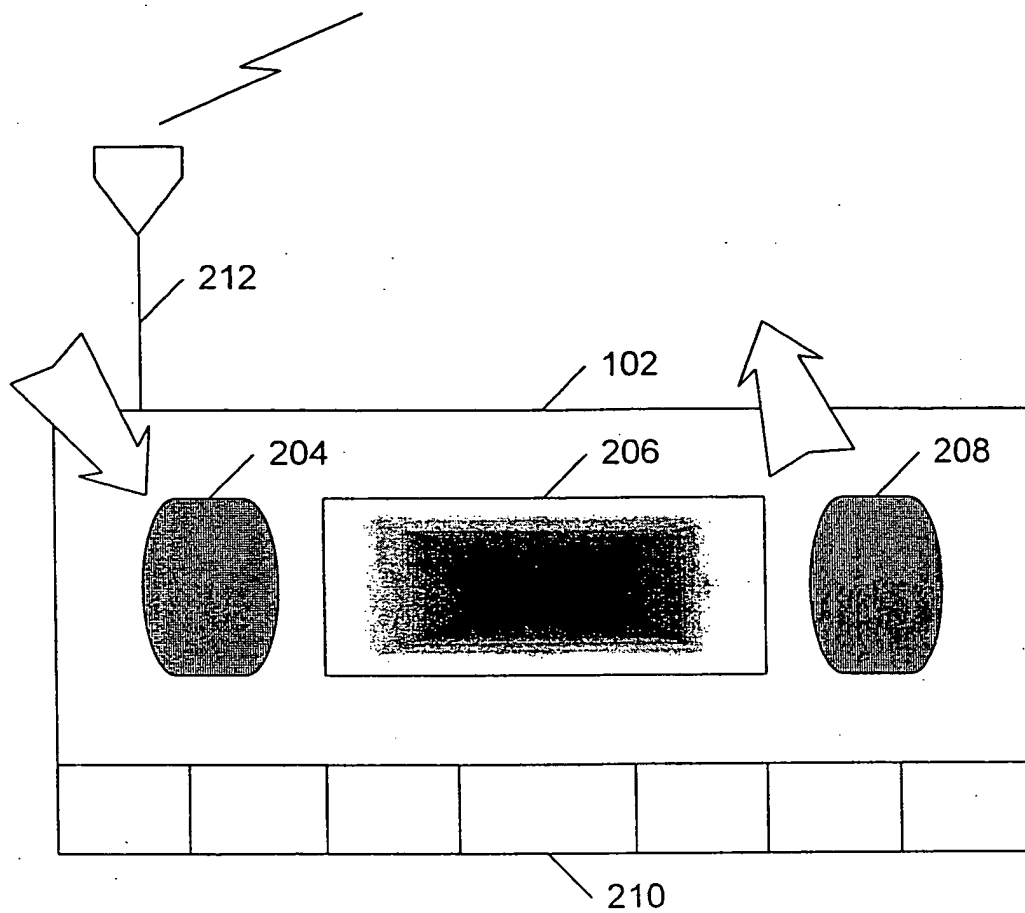


FIG. 2

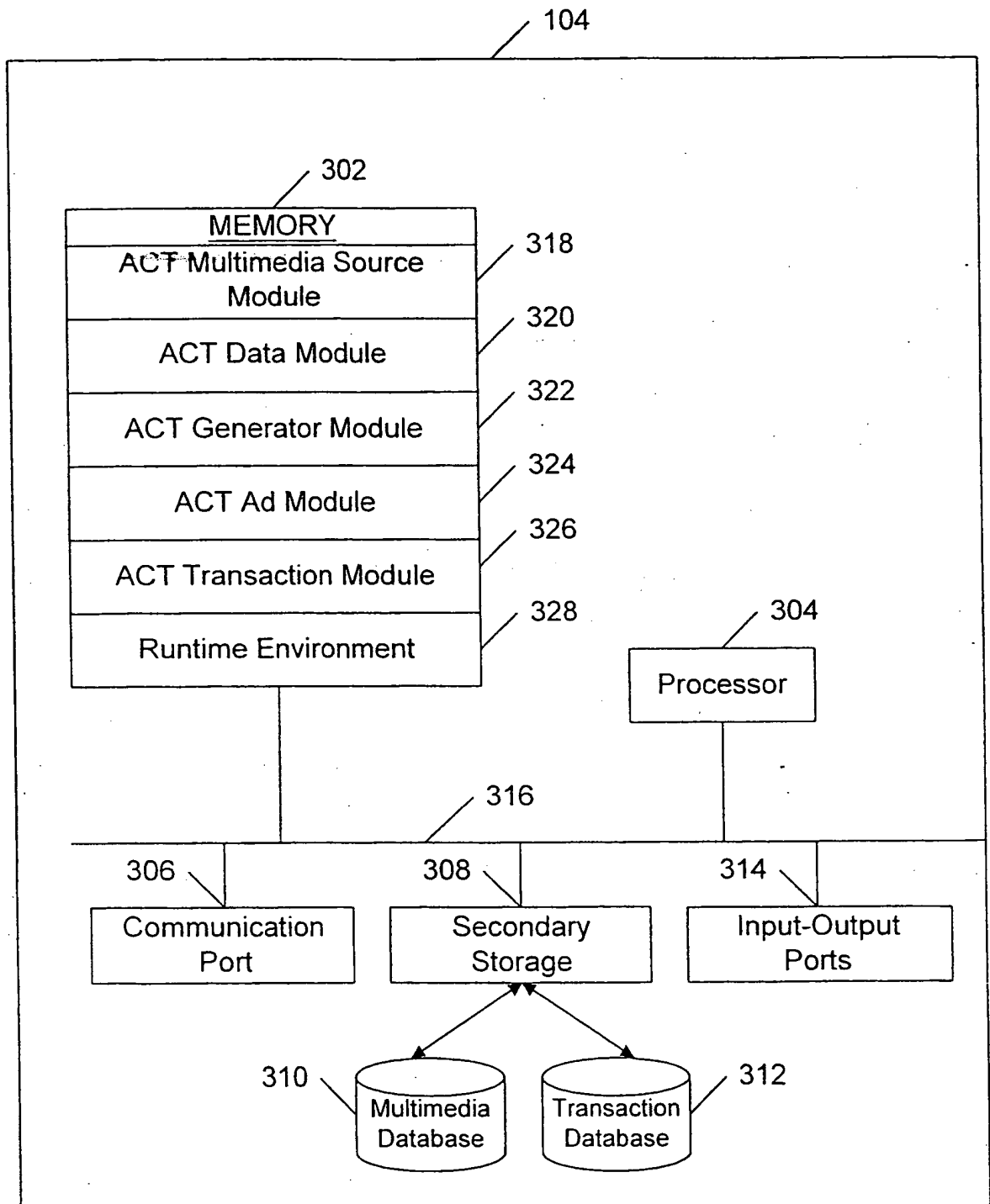
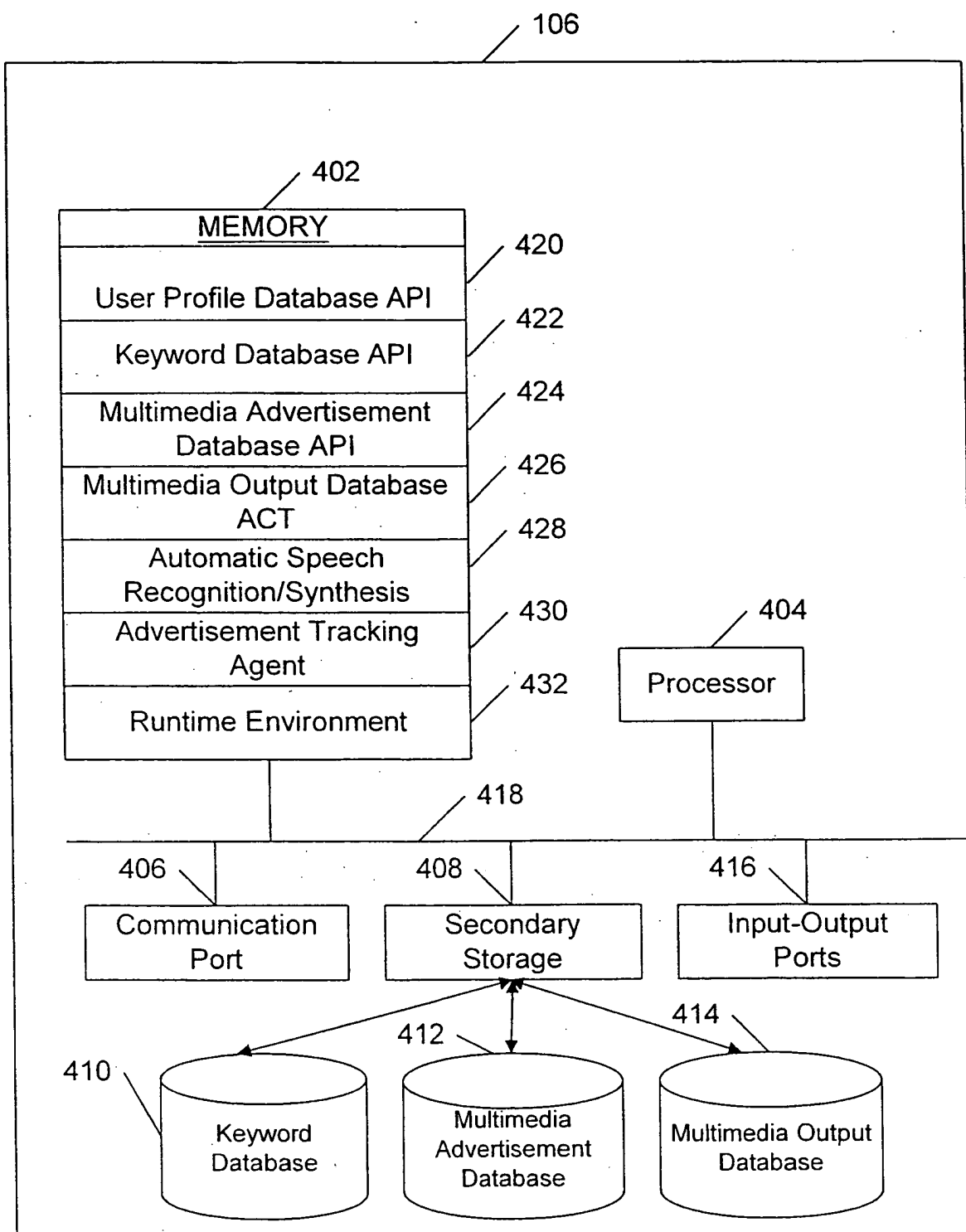


FIG. 3.

**FIG. 4**

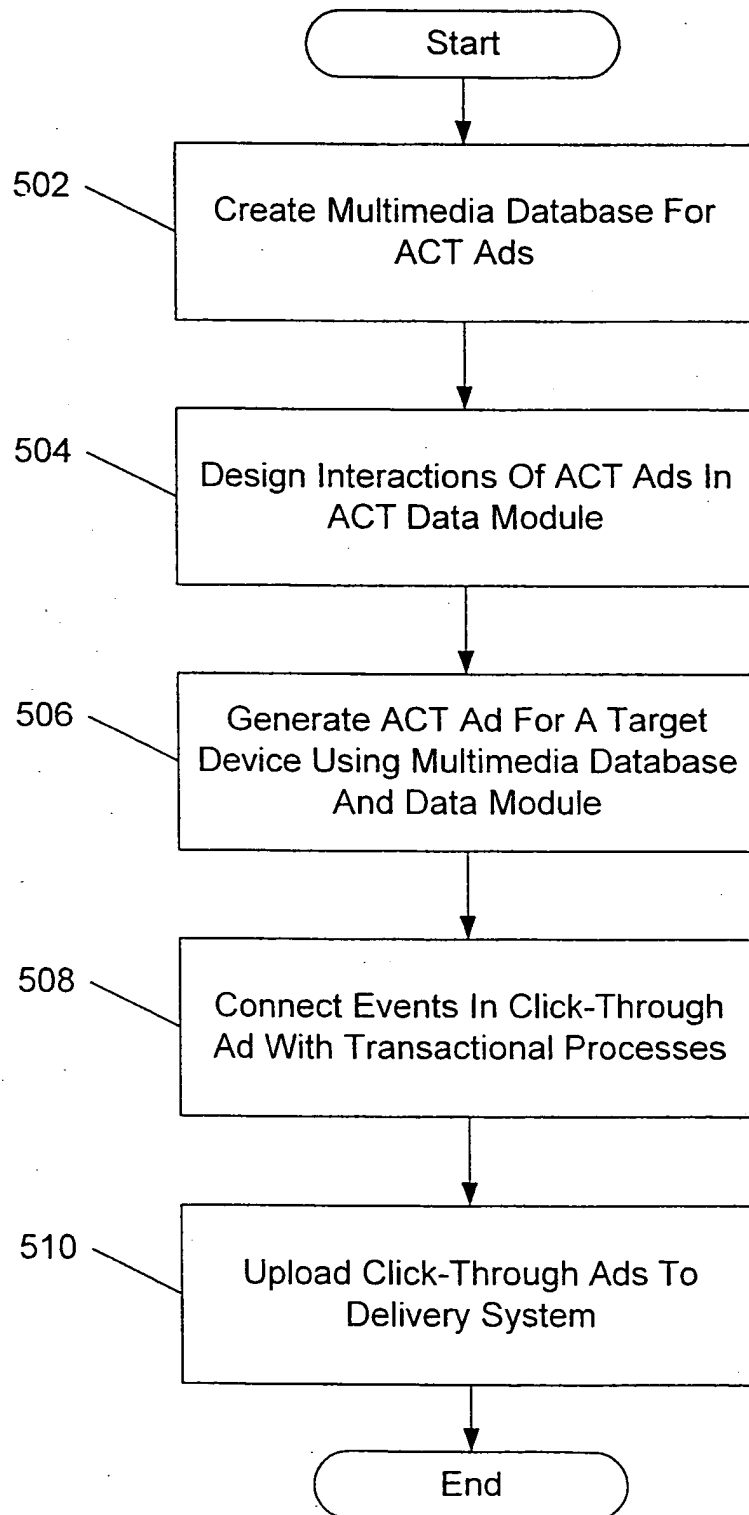


FIG. 5

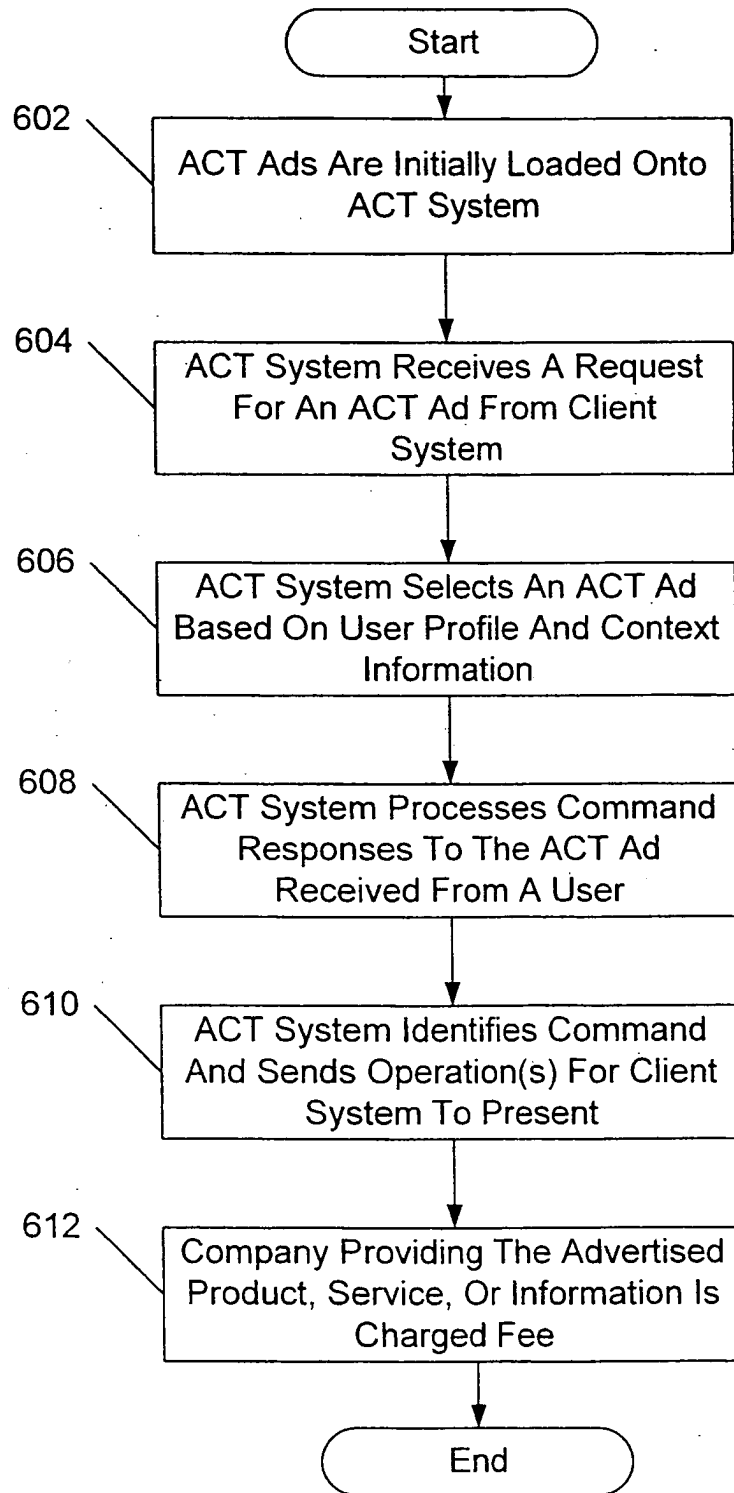


FIG. 6

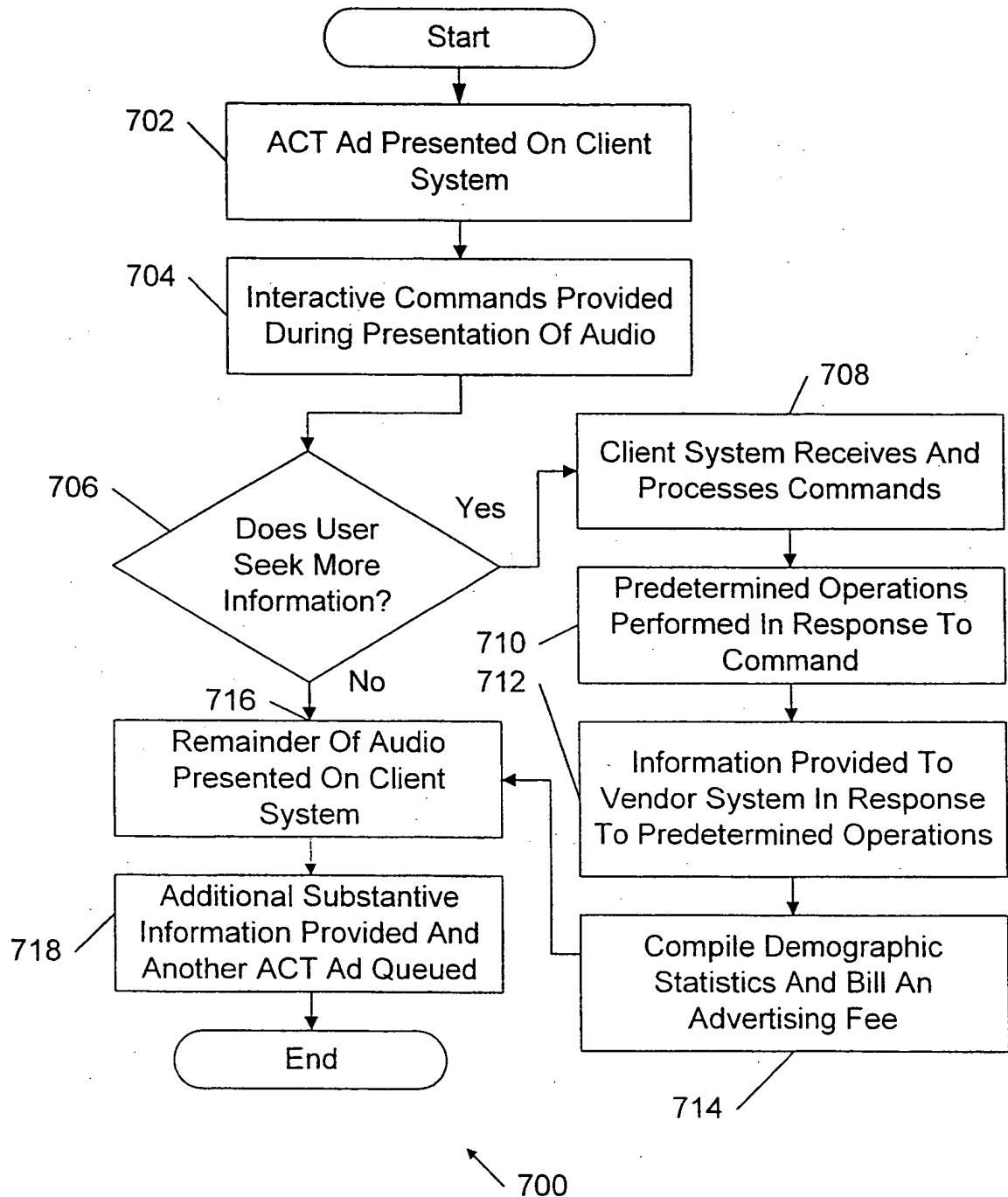


FIG. 7



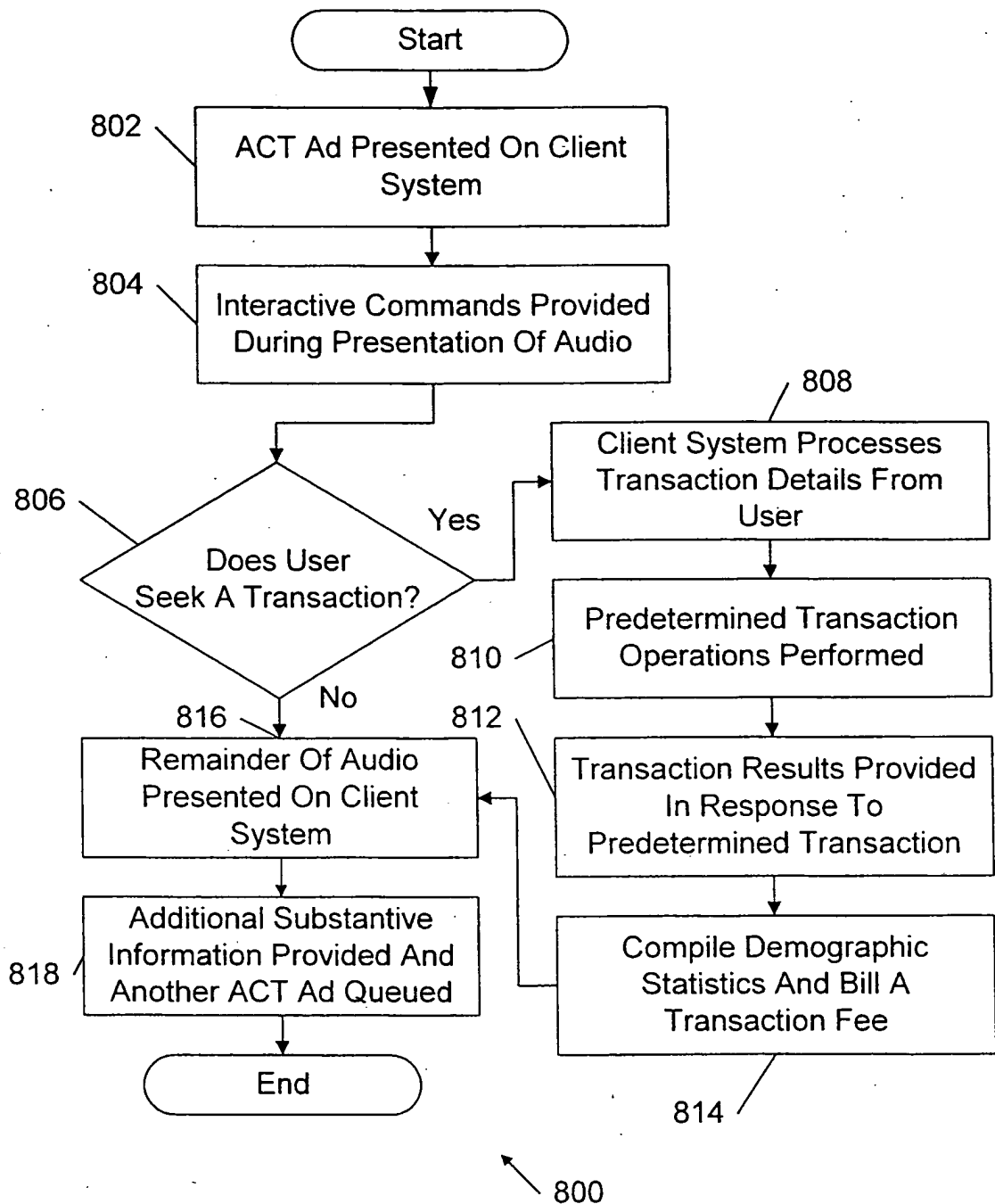


FIG. 8

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/41878

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : Please See Extra Sheet.

US CL : 705/1, 10, 14

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 705/1, 10, 14

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
Microsoft Press COMPUTER DICTIONARYElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
EAST, WEST, PALM Intranet, DIALOG

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y,P	US 6,119,101 A (PECKOVER) 12 September 2000, col. 14, lines 15-54.	1-38
Y	US 5,991,735 A (GERACE) 23 November 1999, col. 5, lines 54-67.	1-38
Y	US 5,948,061 A (MERRIMAN et al) 07 September 1999, col. 5, lines 10-63.	1-38
Y	US 5,848,396 A (GERACE) 08 December 1998, col. 1, lines 15-53.	1-38

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

09 MARCH 2001

Date of mailing of the international search report

04 APR 2001

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231Authorized officer  
STEVE GRAVINI

Peggy Hanoch

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/41878

## A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

G06F 17/40, 17/60, 19/00

**THIS PAGE IS BLANK**